

TECHNIQUE FOR OPTIMIZING BACKOFF FOR A SHARED RESOURCE

Cross-Reference to Related Applications

[0001] This application claims the benefit of United States Provisional Patent Application Serial Number: 60/455,322, entitled "Backoff Substitution on a Shared Communications Channel," filed on March 17, 2003 (Attorney Docket: 680-064us), which is incorporated by reference.

Field of the Invention

[0002] The present invention relates to telecommunications in general, and, more particularly, to local area networks (LAN).

Background of the Invention

[0003] Figure 1 depicts a schematic diagram of network 100 in the prior art. Network 100 operates in accordance with a set of protocols, and comprises shared resource 102 and telecommunication stations 101-1 through 101-K, wherein K is a positive integer. Network 100, for example, can be an Institute of Electrical and Electronics Engineers (IEEE) 802.11 wireless local area network with shared resource 102 acting as the shared-communications channel. Host computers typically use stations 101-1 through 101-K, in an arrangement of one station paired with each host computer, to enable communications between the host computers or between the host computers and other devices, such as printer servers, email servers, file servers, etc. Station 101-3 operates as an "access point" in that it enables other stations within network 100 to communicate with each other by distributing access parameters and other information pertinent to network 100 as needed.

[0004] Stations 101-1 through 101-K contend with each other to use shared resource 102 to communicate. As a result, stations 101-1 through 101-K must use access rules to resolve the contention that occurs. Each of stations 101-1 through 101-K listens for transmissions (*i.e.*, performs carrier sensing) on shared resource 102. Only if station 101-1, as an example, does not detect a transmission for a fixed period of time known as the "interframe space" does station 101-1 attempt to contend for access to shared resource 102.

[0005] In contending for access, station 101-1 then performs a backoff. The backoff consists of waiting a random number of idle time slots, where each slot in this "slot-based backoff" is a fixed time interval. Station 101-1 chooses randomly the number of backoff slots that it must wait between zero and a contention window parameter value. Only after the backoff period can station 101-1 transmit a frame (*i.e.*, a packet).

[0006] During the backoff period, station 101-1 must continuously monitor shared resource 102 to determine whether or not the channel is idle. This means that station 101-1 must keep its receiver powered on during this time.

[0007] Station 101-1's wait for access to shared resource 102 might be interrupted by frame transmissions by other stations. For example, station 101-2 might attempt to transmit before station 101-1 because station 101-2 generates a number of backoff slots that is less than the number generated by station 101-1. In this case, the backoff countdown process halts for station 101-1 and begins again only when shared resource 102 has been idle for another interframe space after the end of the interrupting transmission.

[0008] The advantage of slot-based backoff is that it resolves contention and minimizes the event of multiple stations attempting to transmit at the same time (*i.e.*, a collision event) while maintaining a relatively high level of efficiency in using shared resource 102. Station 101-1's wait for access, however, can be substantial, especially if other, earlier-to-transmit stations have long transmission times. Consequently, the disadvantage of slot-based backoff is that keeping station 101-1's receiver turned on during the wait for access consumes power.

[0009] What is needed is a technique to access a shared resource that is governed by contention-based access without the receiver power consumption disadvantages in the prior art.

Summary of the Invention

[0010] The present invention is a technique that enables access to a shared resource that is governed by contention-based access while reducing receiver power consumption. The illustrative embodiment of the present invention replaces the slot-based backoff in the prior art with an event-based backoff, in which the event-based backoff interval accounts for other stations contending for the shared resource.

[0011] The illustrative embodiment of the present invention determines the backoff interval that a station should use by measuring the average wait time incurred during previous access attempts governed by slot-based backoff before that station finally gained

access to the shared resource. The wait time comprises possibly many slot-based backoff intervals and transmission intervals caused by other stations gaining access to the shared resource first.

[0012] After the event-based backoff, in some embodiments, the station waits a short, slot-based backoff time to minimize the collisions that might occur when multiple event-based backoffs expire simultaneously.

[0013] The duration of the event-based backoff in the illustrative embodiment is known in advance. Therefore, during the backoff interval, the station powers down its receiver in some embodiments and conserves power.

[0014] An illustrative embodiment of the present invention comprises: using a shared resource; and refraining from contending for access to the shared resource for a backoff interval after the last use of the shared resource; wherein the backoff interval is based on at least one previous backoff interval that is associated with at least one previous use of the shared resource.

Brief Description of the Drawings

[0015] Figure 1 depicts a schematic diagram of network 100 in the prior art.

[0016] Figure 2 depicts a schematic diagram of network 200 in accordance with the illustrative embodiment of the present invention.

[0017] Figure 3 depicts a block diagram of the salient components of station 201-*i* in accordance with the illustrative embodiment of the present invention.

[0018] Figure 4 depicts a flowchart of the illustrative embodiment of the present invention.

[0019] Figure 5 depicts a timing diagram of the illustrative embodiment of the present invention.

[0020] Figure 6 depicts a graphical diagram of the relationship between the utilization of shared resource 202 and the degree to which station 201-*i* uses event-based backoff, in accordance with the illustrative embodiment of the present invention.

Detailed Description

[0021] Figure 2 depicts a schematic diagram of network 200 in accordance with the illustrative embodiment of the present invention. Network 200 operates in accordance with a set of protocols (e.g., IEEE 802.11, etc.) and comprises stations 201-1 through 201-*L*,

wherein L is a positive integer, and shared resource 202, interconnected as shown. It will be clear to those skilled in the art how to make and use shared resource 202.

[0022] In some embodiments, station 201-3 operates as an "access point" in that over time it distributes access parameters and other information pertinent to network 200. It will be clear, however, to those skilled in the art, after reading this specification, how to make and use network 200 without an access point.

[0023] Station 201- i , for $i=1$ to L , is capable of receiving data messages from host computer 203- i and transmitting over shared resource 202 data frames comprising the data received from host computer 203- i . Station 201- i is also capable of receiving data frames from shared resource 202 and sending to host computer 203- i data messages comprising data from the data frames. It will be clear to those skilled in the art, after reading this specification, how to make and use station 201- i .

[0024] Figure 3 depicts a block diagram of the salient components of station 201- i in accordance with the illustrative embodiment of the present invention. Station 201- i comprises receiver 301, processor 302, memory 303, and transmitter 304, interconnected as shown. In some embodiments where host computer 203- i is not associated with station 201- i (i.e., as with station 201-3), the path to host computer 203- i is not present.

[0025] Receiver 201 is a circuit that is capable of receiving frames from shared resource 202, in well-known fashion, and of forwarding them to processor 302. It will be clear to those skilled in the art how to make and use receiver 301.

[0026] Processor 302 is a general-purpose processor that is capable of performing the tasks described below and with respect to Figures 4 and 5. It will be clear to those skilled in the art, after reading this specification, how to make and use processor 302.

[0027] Memory 303 is capable of storing programs and data used by processor 302. It will be clear to those skilled in the art how to make and use memory 303.

[0028] Transmitter 304 is a circuit that is capable of receiving frames from processor 302, in well-known fashion, and of transmitting them on shared resource 202. It will be clear to those skilled in the art how to make and use transmitter 304.

[0029] Figure 4 depicts a flowchart of the salient tasks performed by the illustrative embodiment of the present invention. It will be clear to those skilled in the art which tasks depicted in Figure 4 can be performed simultaneously or in a different order than that depicted.

[0030] At task 401, a station constituting network 200 uses shared resource 202 in well-known fashion. For example, an IEEE 802.11 station can use shared resource 202 as a shared-communications channel to transmit a frame of data.

[0031] At task 402, station 201-1, used here for illustrative purposes, needs to access shared resource 202 to transmit a frame, but refrains from contending for access to shared resource 202 for a backoff interval after the last use of shared resource 202. The backoff interval is based on one or more previous backoff intervals associated with one or more previous uses of shared resource 202. Each previous backoff interval can comprise one or more slot-based backoffs, or it can comprise one or more previously-occurring backoffs of the illustrative embodiment (*i.e.*, event-based backoff), or it can comprise both.

[0032] Station 201-1 determines the backoff interval ahead of time, in some embodiments, so that station 201-1 is ready to use the backoff interval when needed.

[0033] As part of determining the backoff interval, station 201-1 measures the average wait time that station 201-1 had incurred during one or more previous access attempts before station 201-1 eventually gained access to shared resource 202. As part of the measuring task, station 201-1 observes the periods of backoff that it had experienced during and associated with N frames actually transmitted.

[0034] During this N -frame training period, in some embodiments, station 201-1 tracks the total number of backoff slots it has generated, s , and the total duration of backoff, d , to determine a value for the ratio d/s . For example, Figure 5 depicts a timing diagram in the scenario wherein N is equal to one, in accordance with the illustrative embodiment of the present invention. Shared resource 202 is in use at the time of transmission 501 when station 201-1 determines that it has to transmit a frame. The parameter d is a measure of the total time interval between i) when station 201-1 first sensed shared resource 202 become idle, at time 502, and ii) when station 201-1 eventually transmits its frame (transmission 511) beginning at time 510. In between time 502 and time 510 are alternating periods of idle (*i.e.*, contention intervals 503, 505, 507, and 509) and busy (*i.e.*, transmissions 504, 506, and 508). Transmissions 504, 506, and 508 correspond to when other stations beat out station 201-1 in gaining access to shared resource 202. The parameter d comprises the waits experienced during i) the interframe spaces and backoff slot periods constituting contention intervals 503, 505, 507, and 509, and ii) the transmission times contributed by earlier-to-transmit stations, represented by transmissions 504, 506, and 508. The parameter s is equal to the total number of slots generated by station 201-1 across contention intervals 503, 505, 507, and 509.

[0035] The value of the ratio d/s becomes the *effective* length of each time slot as experienced by station 201-1 (as opposed to the *defined* length of each time slot). Using the ratio d/s enables compatibility with existing backoff-related parameters, as will be explained later. If station 201-1 is one of few stations (or the only station) on the network and the frame transmission time is short, then the ratio d/s approaches the defined length of each time slot (*i.e.*, is small). If, however, there are many stations on the network or each station's frame transmission time is long or both, then the ratio d/s is large.

[0036] The value for N in the N -frame training period in some embodiments can be set to a large number to improve the accuracy in assessing an overall ratio d/s . Alternatively, N can be set to one (or some small number) and station 201-1 can update the ratio d/s as a moving average. It will be clear to those skilled in the art, after reading this specification, how to set a value for N and how to update d/s .

[0037] Once station 201-1 determines the ratio d/s , it calculates the backoff interval as being equal to:

$$(d/s)*(BW/2) \quad (\text{Eq. 1})$$

wherein BW is the backoff window in slots. BW can be set, for example, to the IEEE 802.11 parameter for minimum contention window (CW_{min}). It will be clear to those skilled in the art, after reading this specification, how to tailor the backoff interval using BW . It will also be clear to those skilled in the art, after reading this specification, how to use the waiting time observed to determine backoff interval in a different way than by using Equation 1. For example, the average waiting time per frame to be transmitted can be used as is or with a percentage scaling factor applied.

[0038] It will be clear to those skilled in the art, after reading this specification, how to make and use other techniques to determine a backoff interval that is based on one or more previous backoff intervals.

[0039] At task 403 in some embodiments, station 201-1 powers down receiver 301 in well-known fashion during at least a portion of the backoff interval. Powering down receiver 301 reduces power consumption, extending battery life where used.

[0040] At task 404 in some embodiments, station 201-1 transmits a frame using shared resource 202 in well-known fashion after refraining for the backoff interval.

[0041] Station 201-1 can repeat the N -frame training period going forward or not. If station 201-1 repeats the N -frame training period going forward, station 201-1 can determine the average waiting time based on either slot-based backoff or the event-based

backoff of the illustrative embodiment. Some or all of stations 201-1 through 201- L can alternate between using slot-based backoff and the event-based backoff of the illustrative embodiment, or continue to use slot-based backoff only, or continue to use solely the event-based backoff of the illustrative embodiment.

[0042] The backoff interval determined, in some embodiments, applies for the next M frames to be transmitted by one or more stations constituting network 200. The relative values of N and M influence the efficiency of the determined backoff interval of the illustrative embodiment versus the reaction time to the changing conditions of shared resource 202. It will be clear to those skilled in the art, after reading this specification, how to set a value for M , both independent of and with respect to N .

[0043] After the event-based backoff, as part of the backoff interval in some embodiments, station 201-1 waits a slot-based backoff period. The slot-based backoff period can be introduced to minimize the collisions that might occur when multiple event-based backoffs expire simultaneously. The slot-based backoff process comprises determining a random wait interval in well-known fashion. The slot-based backoff period can be set relatively short, in the sense that the contention window used is small, relative to the overall backoff interval. In some embodiments, the slot-based backoff period can be a nonzero value only after a collision has already occurred that results in an unsuccessful transmission attempt. This condition can be met in IEEE 802.11, for example, by setting CW_{min} to zero and CW_{max} to a nonzero value.

[0044] In some embodiments, the backoff interval is constrained to be at least equal to interframe periods known in the art (e.g., IEEE 802.11 distributed coordination function interframe space [DIFS], etc.). This constraint can be introduced for reasons of compatibility with other stations that do not use the backoff of the illustrative embodiment.

[0045] The technique used to determine the backoff interval can be located at each station 201- i (e.g., station 201-1 as in the example provided, etc.), or it can be centralized (e.g., at an access point such as station 201-3, etc.). If the technique is located at a centralized point, such as an access point, the centralized point updates and distributes a value representing the backoff interval on an ongoing basis, as well as other relevant information, through management frames such as beacon frames. It will be clear to those skilled in the art how to distribute information through management frames. Station 201- i receives the value and uses it to determine the backoff interval.

[0046] In some embodiments, station 201-1 uses the backoff interval determined in the illustrative embodiment only when in a power save mode that other transmitting

stations (e.g., station 201-3 serving as an access point, etc.) are aware of. When in a power save mode, station 201-1 may power down its receiver, its transmitter, or both. When other stations are aware of station 201-1 being in power save mode, the other stations do not transmit frames to station 201-1 and, as a result, station 201-1 does not receive frames during the time that it has powered down receiver 301. When not in a power save mode, in some embodiments, station 201-1 uses slot-based backoff in well-known fashion rather than the event-based backoff of the illustrative embodiment.

[0047] Figure 6 depicts the relationship between shared resource utilization and the degree to which event-based backoff is used, in accordance with the illustrative embodiment of the present invention. Shared resource utilization refers to the percentage of time that shared resource 202 is busy (*i.e.*, utilized). When shared resource 202 is not busy (*i.e.*, is idle), it can be because of several reasons. One reason is that when shared resource 202 becomes idle, it stays idle for at least the interframe space period plus the backoff interval (*i.e.*, slot-based or event-based) in effect.

[0048] Furthermore, since each event-based backoff interval is typically longer than each of the individual slot-based backoff intervals, there is more of a likelihood that shared resource 202 becomes available, but no station is able to transmit shortly thereafter. The more event-based backoff is used over slot-based backoff or 2) the longer the backoff interval, the less shared resource 202 is utilized. The reverse is also true in that 1) the less event-based backoff is used or 2) the shorter the backoff interval, the more shared resource 202 is utilized. This relationship is why, in some embodiments, a scaling factor is used to adjust the event-based backoff interval. The scaling factor can be used to tune the utilization of shared resource 202 by adjusting the degree to which the illustrative embodiment is used, even if the receiver power consumption changes as well. It will be clear to those skilled in the art, after reading this specification, how to tune the degree to which event-based backoff is used.

[0049] It is to be understood that the above-described embodiments are merely illustrative of the present invention and that many variations of the above-described embodiments can be devised by those skilled in the art without departing from the scope of the invention. It is therefore intended that such variations be included within the scope of the following claims and their equivalents.

[0050] What is claimed is: